National Institute Economic Review (2022), 259: 1, 20–30
doi:10.1017/nie.2022.2

NIER CAMBRIDGE UNIVERSITY PRESS

ARTICLE

THINKING ABOUT SOCIAL NORMS

Ken Binmore*

Economics Department, Bristol University, Bristol, United Kingdom
*Corresponding author. Email: k.binmore@ucl.ac.uk

Abstract
This article argues that it is a waste of time seeking to treat populists as examples of *homo economicus* when seeking to persuade them that the conspiracy theories to which they subscribe are big lies. But it does not follow that *homo economicus* is worthless in this context. He has a role in explaining the evolution of the social norms whose violation is the root cause of the rise of populist movements. Such an approach requires a willingness to entertain both proximate and ultimate explanations of human behaviour simultaneously.

Keywords: proximate; ultimate; *homo economicus*; social norm; convention; reciprocity; folk theorem

JEL codes: A13; B00; B52; C70

1. Introduction

This article argues that it is a waste of time seeking to treat populists as examples of *homo economicus* when seeking to persuade them that the conspiracy theories to which they subscribe are big lies. But it does not follow that *homo economicus* is worthless in this context. He has a role in explaining the evolution of the social norms whose violation is the root cause of the rise of populist movements.

Attempts to put across the ideas addressed in this article to audiences for whom everything is new typically get mired down in a mess of detail. What is the Prisoners' Dilemma? Who is John Nash? How does the strategy *tit-for-tat* work? What is Pareto efficiency? Does so-and-so really deserve credit for such-and-such? Why is this or that argument fallacious? Why does some definition or another need to be qualified to take account of pathological cases? This article takes for granted an audience that is tired of having such matters explained. Let me recommend my book *Natural Justice* to readers for whom this assumption is mistaken.

2. Homo economicus

Nobody denies that the *homo economicus* paradigm works well in describing human behaviour in classical markets, but critics commonly argue that it needs to be replaced by something more inclusive to capture how we behave in more general social situations.

One of the objectives of this article is to argue that such critics often make the mistake of thinking that *homo economicus* can only be deployed in trying to model the social world as a kind of market, but the success of *homo economicus* as a tool in evolutionary biology should surely put paid to this parochial attitude. However, the area on which this article focuses is the use of *homo economicus* by game theorists to model social norms that depend on the kind of long-term cooperative relationships that the Chicago model of a one-shot market excludes.

It is natural to respond that the superhuman rationality attributed to *homo economicus* is utterly unrealistic, but the same is true when *homo economicus* is used to model ants in biological applications, although nobody thinks that ants are superhumanly rational. So why does it work so well when applied to ants? (e.g. Pollock *et al.*, 2012.)

© The Author(s), 2022. Published by Cambridge University Press on behalf of National Institute Economic Review.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

2.1. Ultimate and proximate explanations

The answer is to be found in the four modes of explanation famously identified by the biologist Tinbergen (1963). The two that are relevant here are ultimate explanations and proximate explanations.

For example, why do songbirds sing in the early spring? A proximate explanation in fundamental terms would be very complicated. This molecule knocked against that molecule. This neuron fired in response to the firing of those neurons. That chemical reaction was catalysed by this enzyme.

An ultimate explanation can be very much simpler. Birds sing in the early spring to advertise their presence to potential mates and possible rivals. They neither know nor care that this is why they sing. They just do what they do. Birds in the past that failed to do so lost out in the evolutionary competition to replicate their traits, and so only birds that sing in the early spring survive today.

It is pointless to ask which of these explanations is correct. Both are correct. One supplies the answer to *why* the behaviour exists. The other to *how* the behaviour is implemented. This article argues that we should allow ourselves the same freedom to entertain multiple explanations operating at different levels when thinking about social norms.

What function does a social norm serve? Why does it survive? For such ultimate questions, the *homo economicus* paradigm can be useful, although nothing says that other models may not be better in particular cases. How is a social norm implemented? When asking such a proximate question, we have yet another choice to make between multiple levels of explanation. Do we appeal to physics? Classical or quantum? Or should we go up a level to chemistry or biology? None of these proposals is practical. The lowest level at which we might be able to write a tractable model is psychological. In particular, we can ask how people explain to themselves why they carry out the actions necessary to sustain a social norm.

The answer to such a proximate question is unlikely to coincide with the ultimate explanation of a social norm because people need a personal reason to explain their own behaviour. Why do I get angry when others deviate from a particular social norm? Why do I feel the urge to punish deviants with my disapproval or worse? Is it for religious reasons? Am I bound by some code of honour? Is it because I feel the urge to copy what I see others doing? Am I fearful of how I will be treated if I do not fall in line? In answering such proximate questions, *homo economicus* is unlikely to be of any use at all.

The plan in the rest of this article is to outline how game theorists use the *homo economicus* paradigm to provide an ultimate explanation of what social norms are for, and how they work. This philosophical section is intended to explain why accepting such an ultimate explanation does not preclude the various attempts at proximate explanations to which reformers commonly appeal with varying success. However, it has to be admitted that most of the commentaries on Hindricks and Guala (2015) when they propose a similar multi-layer attitude to the study of social norms are not very encouraging.

3. De gustibus non est disputandum

There is juvenile joke about getting into a cinema without a ticket by walking in backwards. Some behavioural economists unconsciously practice a similar trick by trying to model social *norms* as social *preferences*, thereby keeping *homo economics* alive in a proximate context while simultaneously proclaiming his death.

Homo economicus is said to be dead because experiments show that people are not selfish, but nothing in the *homo economics* paradigm says otherwise. He certainly maximises his own utility function, but the idea that he cannot derive utility from helping others is a mischievous invention. Neoclassical textbooks say nothing at all about how selfish or unselfish we should assume *homo economicus* to be. Social preferences are no less neoclassical than selfish preferences. Indeed, I complain in Binmore (2010, 2020) that the reason an exclusive focus on social preferences is not very successful in proximate studies of social norms is that it is *too* neoclassical.

In the game-theoretic approach to social norms outlined in this article, the social preferences of the players are just one of the factors that need to be taken into account when determining what payoffs need to be written into the game of life to be considered next.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

4. Game of life

In tribute to the pioneering work of Hume (1739), philosophers often refer to social norms as conventions. My own tribute to the great David Hume in this article consists of an Appendix that lists a series of quotes from his work that show that he already knew much of what I have to say here. Among the many others who have rediscovered or have been inspired by his work, I shall mention only Bicchieri (2006, 2012), Lewis (1969), Schelling (1960) and Young (2001).

As I say in Binmore (2005), social norms or conventions are many and various. They range from the arcane table manners we employ at formal dinner parties to the significance we attach to the green pieces of paper we carry around in our wallets bearing pictures of past presidents. From the rules that govern how we drive our cars in heavy traffic to the meaning of the words in the language we speak. From dietary and sexual taboos to the standards of integrity and truthfulness expected of honourable folk. From the vagaries of fashion to the criteria that we fondly suppose secure ownership of our possessions. From the amount that we think it appropriate to tip in restaurants to the circumstances under which we are ready to submit ourselves to the authority of others.

The next section explains how all such social norms or conventions can be regarded as equilibrium selection devices in the game of life that we play together. But two things need to be made clear from the beginning.

4.1. Natural laws?

Firstly, to follow this modelling strategy, the rules of the game of life must be conceived of as being beyond our control, like the laws of physics, or the facts of biology or geography. This does not imply that the game of life cannot change—only that its players cannot change it.

It is important that rhetoric about invented natural laws that are said to govern our behaviour must be put aside. As Hume explains, it is natural that societies have conventions that govern behaviour within that society, but not that biological evolution has programmed us with some particular convention. Similarly, it is natural that we speak a language, but not that biology has programmed us with some particular language. What philosophers call natural laws—or economists call institutions—correspond to the rules that need to be followed in sustaining a particular equilibrium in the game of life.

4.2. One-shot games?

Secondly, we must look beyond the simple one-shot coordination games used by Schelling (1960) and then Lewis (1969) to illustrate how conventions operate to select a Nash equilibrium from a game with many Nash equilibria.[1] It is certainly helpful to illustrate the idea of a social norm by explaining how the convention of driving on the right selects one of the two pure Nash equilibria of the Driving Game, but we need to consider much more complicated models of the game of life if we are to capture how *reciprocity* can sustain the cooperative social norms on which the success of our species depends. Here again, it is necessary to repeat that we are not a species of social insect programmed with only one way to cooperate. Insofar as our biological programming is relevant, it lies in our being equipped with programs to sustain whatever social norm cultural evolution may have installed in whatever society we inhabit.

It cannot be emphasised enough that although both biological and cultural evolution matter, the kind of evolution we are talking about in discussing the origins of particular social norms is *cultural* evolution and not *biological* evolution. One cannot use models of biological evolution—no matter how successful they may be in biology—and proceed as though they will also work in a cultural context. In particular, it is not enough to replace genes with memes and carry on as though everything biologists have discovered about how biological evolution also works when applied to cultural evolution.

[1]A Nash equilibrium is a strategy profile—one for each player—in which each player's strategy is optimal given the strategy choices of the other players. It is unusual for a game to have only one equilibrium.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

5. Modelling social norms

The basic idea to be explored here is that social norms exist as equilibrium selection devices in the human game of life. We need to restrict our attention to Nash equilibria, because only Nash equilibria satisfy the minimal requirements for evolutionary stability. This seems a crippling requirement when attention is restricted to one-shot games like the Prisoners' Dilemma, but when we expand the scope of our inquiry to indefinitely repeated games, we end up with a theory capable of explaining how efficient cooperation may have evolved in the human species.

5.1. Reciprocity

The idea that reciprocity lies at the root of human cooperative norms goes back at least as far as the Chinese sage we call Confucius.

But we cannot model the idea that I will scratch your back if you will scratch mine in a one-shot game —a game that is imagined to take place at a single instant in time. Nor can reciprocity be modelled in a repeated game with a fixed horizon because cooperative behaviour in the final repetition cannot be rewarded or punished in the future. So any attempt to sustain cooperation through reciprocity will unravel as the implications are brought forward in time. The minimal setting for studying reciprocity is therefore a repeated game with an indefinite time horizon.

The simplest example is the case of the Prisoners' Dilemma played repeatedly, with a fixed probability that the current play of the game will be the last.[2]

The strategy *tit-for-tat* provides an example of how reciprocal cooperation can be sustained in such a game.[3]

Nowadays we would say that if both players use the strategy *tit-for-tat*, the result will be that cooperation is sustained as a Nash equilibrium if the continuation probability is sufficiently large. That is to say, neither player will have an incentive to deviate from *tit-for-tat*, unless the other player deviates first. It is necessary for evolutionary stability—but by no means sufficient—that the strategies in use in a game make up an equilibrium in the sense of Nash.[4]

5.2. Folk theorem

The general result is called the folk theorem in game theory.[5] It works if we start with any one-shot game whatever, and not just the Prisoners' Dilemma.

There will usually be cooperative outcomes of the underlying game that the players would like to achieve, but are unavailable without a legal system to enforce contracts, because they are not Nash equilibrium outcomes—any agreement to cooperate on such an outcome will be unstable because at least one of the players will have an incentive to cheat on the agreement. But if the players care about the future sufficiently, then an approximation to any outcome in the underlying game on which the players might agree if they could write enforceable contracts is available as a Nash equilibrium outcome in the indefinitely repeated game.

[2]Chapter 4 of Binmore (2020) explains the Prisoners' Dilemma, and refutes various fallacious analyses. It also explains the folk theorem to be discussed next without any mathematics.

[3]Tit-for-tat is the strategy—famously popularised by Axelrod (1984)—of beginning by cooperating in the repeated Prisoner's Dilemma, and then copying whatever the opponent did in the last round.

[4]Evolutionary stability is to be interpreted in its general sense here. The formal definition of an ESS by Maynard Smith (1982) is only adequate in the two-strategy, symmetric games for which it was designed.

[5]The idea originates with Hume (1739), but was rediscovered independently by a number of game theorists in the 1950s after Nash defined his equilibrium concept—which is why it is not attributed to any particular person. It was rediscovered 20 years later by the biologist Trivers (1971), who had the good idea of calling the principle *reciprocal altruism.* The final rediscovery by Axelrod (1984) came 10 years or so after that.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

My own opinion is that the folk theorem just stated is of paramount importance for political theory. It would be even more important if it were not subject to an important proviso—the players' history of past play must be an open book. This hardly seems worth mentioning in the two-player case, but it is a major handicap when the folk theorem is applied to a whole society. How much of the folk theorem survives when the players have secrets from each other? Who committed that crime? Who is the ringleader of that rebellious group? We have only examples to guide our intuition in such cases of imperfect information, but perhaps game theory will be able to do better someday.

5.3. Markets

The folk theorem also applies when the underlying game is a classical market. It tells us that the range of possible equilibrium outcomes is very wide. In particular, when the future matters, stable long-term relationships can evolve. We can then no longer safely assume that the price at which goods are traded can be found by equating current supply and demand, as in the Chicago wheat market.

The Marseille daily fish market is a good counter-example. Only some of the market is competitive. The remaining buyers and sellers insure each other against price fluctuations by establishing long-term relationships in which they trade at non-competitive prices (Kirman and Vriend, 2000). But the folk theorem tells us that we do not need to depart from the *homo economicus* paradigm to explain such phenomena.

5.4. More on homo economicus

There are two interpretations of a Nash equilibrium. The success of game theory in evolutionary biology can be traced to the possibility of passing back and forward between the two interpretations.

The first interpretation is as a necessary condition for the rational solution of a game in which the players are traditional representatives of the species *homo economicus.* The second interpretation treats the game's payoffs as measuring the fitness of players who may not be rational at all, but whose play is subject to an adjustment process that favours the more fit at the expense of the less fit. Such a process can only stop operating when it gets to a Nash equilibrium, because only then will all the players be as fit as it is possible to be given the behaviour of the other players.

That is to say, evolution—whether cultural or biological—will tend to generate the same strategies that would be chosen by fully rational examples of *homo economicus.* We can therefore hope to predict the long-term operation of proximate adaptive processes simply by asking the ultimate question: How would examples of *homo economicus* play this game?[6]

But it does not follow that we thereby commit ourselves to the absurd notion that real players are superhumanly rational. I have met several very successful stockbrokers who I would not even trust to find their way around a goldfish bowl, let alone to equate supply and demand in a market.

5.5. Evolution of efficient norms

Axelrod's (1984) evolutionary simulations generated Pareto-efficient outcomes, but there is a co-evolutionary argument for why we should expect to observe[7] social norms that select Pareto-efficient equilibria more generally.

Suppose that many identical small societies are operating one of two social norms, *busy* and *idle.* If *busy* makes each member of a society that operates it biologically fitter than the corresponding member

[6]Perhaps historians of intellectual thought will one day connect this development with the Chicago writings of Alchian (1950) on the one hand, and Friedman (1953) on the other.

[7]Co-evolutionary arguments assume that cultural and biological evolution can operate in tandem, as presumably in the evolution of language.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

of a society that operates *idle*, then here is an argument which says that *busy* will eventually come to predominate.

To say that a citizen is biologically fitter means that the citizen has a larger number of children on average. Societies operating the social norm *busy* will therefore grow faster. Assuming societies cope with population growth by splitting off colonies which inherit the social norms of the parent society, we will then eventually observe large numbers of copies of societies operating the social norm *busy* compared with those operating the norm *idle.*

Part of our success as a species of social animal presumably lies in the fact that we are not biologically programmed to operate any particular version of the *busy* social norm. A society is then free to adapt its choice of social norm to whatever version of the game of life it finds itself playing.

5.6. Evolutionary drift

Axelrod's (1984) enthusiasm for the particular strategy *tit-for-tat* has persuaded some political scientists that there is nothing more they need to know about how reciprocity works, but this is a bad mistake. The folk theorem tells us that indefinitely repeated games generically have an infinite number of efficient equilibria, and it is necessary to come up with extra evolutionary arguments if it is to be argued that one of these is to be favoured over another.[8]

In fact, evolutionary simulations in the repeated Prisoners' Dilemma—including Axelrod's simulations—typically end up drifting among a whole set of efficient equilibria.

The immediate point is that it would be a mistake in the general case to wear a *homo economicus* hat while seeking a precise prediction of which social norm evolution will select. It is seldom a question for rational debate. One might as well ask why a shoal of fish or a murmuration of starlings suddenly ceases to move in one direction, and then moves in another. The answer is that lot of independent random factors sometimes happen to align.

5.7. Multi-player games

To focus on *tit-for-tat* at the expense of its many rivals as the possible end-product of an evolutionary process is particularly misplaced in the case of games with more than two players. One reason is that it need not be the victims of a deviation from a social norm who would punish the deviant. In real life, such victims will sometimes be crippled or dead! However, the natural example of how a tyranny can survive as a social norm is postponed until after the next section, in which the subject of punishment gets a more thorough airing.

A second reason that multi-player games present greater difficulties is even more serious. In such cases, the weakness of Nash equilibrium as a means of modelling evolutionary stability becomes evident. With many players, it is necessary to confront the fact that Nash equilibria are almost never proof against coordinated deviations by whole coalitions of players. Why then do we not restrict attention to strategy profiles that are proof against all such potential deviations? Because such strategy profiles usually do not exist.

6. Punishment

The reason that an indefinitely repeated game of life can sustain efficient Nash equilibria is that deviations from the equilibrium strategy can be punished in the future. For example, if Alice deviates with a *tat,* Bob can respond in the next period with a *tit.*

[8]Binmore (2005) offers such reasons for fairness norms in cases where the underlying game admits a bargaining interpretation. Binmore (2001) gives arguments why *tit-for-tat* is not an adequate representation of reciprocity by itself. Binmore and Samuelson (1994) discuss evolutionary drift more generally.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

Such talk of punishment evokes visions of torture chambers and the gibbet. Countless heretics have suffered in this way, but nearly all the punishments necessary to sustain an established social norm are administered without either the punishers or the victim even being aware that a punishment has taken place. No stick is commonly flourished. What happens most of the time is that the carrot is withdrawn a tiny bit. Shoulders are turned slightly away. Greetings are imperceptibly gruffer. Eyes wander elsewhere. These are all warnings that deviants ignore at their peril.

I was particularly delighted to find anthropological accounts describing higher stages of punishment observed among hunter-gatherer societies, because they mirror so accurately similar phenomena that the academic world uses to keep rogue thinkers in line. First, there is laughter. If this does not work—and who likes being laughed at—the next stage is boycotting. Nobody talks to the offending party or refers to their research. Only the final stage is maximal: persistent offenders are expelled from the group or are unable to get their work published.

6.1. Perfect equilibria

A perfect equilibrium is a Nash equilibrium that remains a Nash equilibrium when attention is restricted to all subgames of the original game—whether or not these subgames will actually get to be played.[9]

We need the concept of a perfect equilibrium to answer the question of why people required to punish a deviation in a Nash equilibrium would carry through with this duty if the deviation were to occur. What is their incentive? Why not leave this possibly unpleasant duty to someone else? The answer is that the folk theorem works not only for Nash equilibria but for perfect equilibria as well.

6.2. Why punish?

If an out-of-equilibrium deviation forces us into a subgame off the equilibrium path, it remains optimal in a perfect equilibrium to stick to your original strategy, provided everybody else does. In particular, if your strategy tells you to punish the deviant at some apparent cost to yourself, it will nevertheless be optimal for you to carry through on the punishment. If you deviate by trying to escape your duty to punish, you will take us to yet another subgame, in which it is optimal for some other player or players to punish you for your dereliction of duty. If they fail to discharge their duty, we go to yet another subgame —and so on forever.

The fact that an infinite regress is involved does not imply that any players need to do anything infinite. All that is needed to implement a perfect equilibrium is that players can distinguish between insiders and outsiders. The insider rules of conduct include the provision that outsiders are punished. Apparent insiders who fail to punish an outsider reveal themselves as outsiders.

6.3. Authority

David Hume pointed out long ago that there is an apparent paradox in the survival of tyrannies because power ultimately rests in the collective hands of the oppressed. His answer is that it is only by convention that somebody in authority is obeyed. In game theory terms, even the most tyrannical leader can do no more in the long run than serve as a coordination device, whose ultimate role is to choose between the available perfect equilibria.

Alice in Wonderland obeys the orders of the Queen of Hearts because she anticipates that otherwise the Queen will order Bob to cut off her head. Bob would obey this order because he anticipates that otherwise the Queen will order Carol to cut off his head. Carol would obey the order to punish Bob

[9]Reinhard Selten was the first to look at perfect equilibria. Some confusion is possible because he changed the definition offered in the text to something more general in his later work. When it is necessary to distinguish between the two definitions, the original definition is called *subgame perfect* and the later definition is called *trembling-hand perfect.* However, neither kind of perfect equilibrium is any better than Nash equilibrium at resisting deviations by coalitions.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

because she fears that someone else would cut off her head. Since this someone else might be Alice, we are looking at a spiral of self-confirming beliefs, which may seem too fragile to support anything solid. It is true that the beliefs go round in a circle, but the folk theorem shows that this fragility is an illusion, since the behaviour generated by the beliefs holds together as a perfect equilibrium.

The role of Donald Trump since he lost the 2020 presidential election is a good example. He commands no army or navy, but nevertheless directs the immediate future of the Republican party like a puppet master. He does so merely by pointing at the perfect equilibrium he wants to see implemented. Those who deviate from this equilibrium are punished. Those who fail to punish the deviants become deviants themselves.

7. The role of the emotions

This section ceases to address ultimate explanations and turns its attention instead to proximate explanations. How does it feel to be a player in a game operating a particular social norm? *Homo economicus* has no feelings at all, but we are no longer talking about him. We have now to attempt the much harder task of talking about ourselves.

Emotions were once dismissed as irrational urges left over from our evolutionary history. The socially aroused emotions associated with pride, envy, and anger are still counted among the seven deadly sins. But if these emotions are as self-destructive as tradition holds, how come they have not been evolved away?

It is now widely accepted that tradition is plain wrong in seeing no useful role for emotions. I think their social role evolved to help police primaeval social norms, and they remain useful to us for this purpose.

In this way, it is possible to sustain efficient equilibria in repeated games without any of the players even being aware that they are playing a game. How else would chimpanzees be capable of sustaining high levels of reciprocal altruism? How would humans be capable of the same feat if we always had to spend half an hour or more calculating what to do before taking any action? Some of our thinking in these situations must surely be hardwired, and perhaps getting into an emotional state is simply how it feels when our autopilot takes over the controls.

7.1. Anger

Sexual norms are traditionally maintained by our being habituated to respond with disgust at certain practices, although the capacity to feel disgusted presumably first evolved for hygienic reasons.

I suspect that biological evolution has similarly co-opted anger to play a policing role in sustaining social norms. That is to say, its ultimate social role is to promote the punishment of deviants when they fail to honour an established social norm.

For example, the prototypical scenario for the expression of social anger arises when Adam treats Eve unfairly. In her anger at his unjust treatment, she then feels an urge to inflict some harm on Adam. Adam, therefore, takes care to keep his acquisitive urges under control lest he provokes her into taking vengeance. Even when the act of taking vengeance may be personally very costly to Eve, she may still feel compelled to do so because of the fear of social exclusion that may result from her failing to do what her friends and relatives regard as her social duty.

It is not hard to persuade people of such scenarios in the case of the fairness norms studied in my *Natural Justice*, but I want to put the same case for less attractive social norms: for example, the social norms that determine authority or status. The suggestion is that anger is provoked whenever *any* established social norm of sufficient importance is dishonoured.

We do not need to assume that the folk rendered angry for this ultimate reason are even aware that their earlier behaviour is mediated by a social norm. They will typically indignantly deny that they are

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

acting no more rationally than a fish in a shoal, and offer rationalisations for their anger that makes little real sense, but it does not follow that such rationalisations can be ignored if the aim is to avoid conflict.

What have social scientists to say about such proximate problems? Very little it seems. We are all appalled by the Trump phenomenon. We agree on our diagnosis, but not on a remedy. We cannot even agree on a model that we could use to evaluate possible remedies. The best we can seemingly do is to rule out certain ways of thinking about Trumpery that do not hold together, as I attempt with a few inadequate remarks in the next section.[10]

8. Trumpery

Basically, what I have to say is only that Trumpery and other kinds of modern populism are merely symptoms of a more fundamental disease that will not go away if we treat only the symptoms.

8.1. Disease?

The disease is that our social contract—the set of social norms that govern our lives—has shifted, partly through technological and other changes in our underlying game of life, and partly through cultural drift. One result is that many of us, especially blue-collar workers without a college education, feel alienated.

In ultimate terms, the response of such folk reveals that human biological programming is ill-adapted to rapidly changing social norms, so they behave as though experiencing a deviation from a fixed social norm. They are programmed to feel anger in such circumstances in order that they be motivated to punish both the deviants, and those who make themselves deviants by not joining them in punishing those perceived as deviants. They would laugh at this ultimate explanation, but they nevertheless need an explanation of some kind to rationalise their anger. So they look around for deviants to blame. Is there a conspiracy of Washington insiders to cheat us of our rightful deserts? Are Latino immigrants stealing our jobs? Does one need to be WOKE to get ahead these days? Is it a crime to be a WASP? And so on.

8.2. Big lie?

It is a waste of time to attempt to confront such angry folk with the objective facts. For example, there is no point in explaining why it is a lie that the 2020 presidential election was stolen. As with the theologian who would not look through Galileo's telescope lest he sees the supposedly non-existent moons of Jupiter, the facts are irrelevant. Similarly, for the creationist grandparents who had reburied the magnificent ichthyosaur fossil, an English guy found last year while digging in his garden.

We are tempted to react to such stories as CNN reacts to Trump's big lie. We treat the proponents of such easily refuted beliefs as though they were examples of *homo economicus* whose mind could be changed by exposure to the facts of the matter. But populist big lies are not to be classified as the kind of beliefs we attribute to *homo economicus.* They are actually rallying calls, whose function is to provide a coordinating focal point for folk whose autopilots have been made angry at a historic social norm apparently being violated.

Insofar as economic theory is relevant here it lies in applying a kind of reversal of the neoclassical theory of revealed belief and preference. Instead of deducing the beliefs and preferences of agents assumed to be rational from their observed behaviour, potential insurgents summarise how their collective power could be brought to bear by stating a set of invented beliefs that would generate the necessary behaviour among rational agents if they were believed.[11]

[10]The dictionary definition of *trumpery* is something attractive but worthless.

[11]The pragmatist Peirce observes: To be deliberately and thoroughly prepared to shape one's conduct into conformity with a proposition is neither more or less than the state of mind called Believing that proposition.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

8.3. Stability at threat

Once the ultimate purpose of a populist lie has been understood, it becomes easy to understand why populists do not choose to investigate its literal truth. But the Republican senators who cynically endorse Trump's big lie are playing with fire because the kind of social norms that can be modelled as perfect equilibria are only stable in the face of individual deviations. Only in very exceptional circumstances can equilibria be found that are stable against the type of closely coordinated deviations by large coalitions that the creation of the internet has made possible.

The immediate disease that needs to be treated is the shift in the social contract that is the root cause of the anger that threatens—not only the long-standing American consensus—but the consensus in my own country and elsewhere. Agonising about the symptoms will get us nowhere. The long-term problem is even more serious. The ultimate stability of our societies has depended in the past on our relying on social norms that implement Nash equilibria that are proof against attack by uncoordinated deviants. But the internet has let a coordination genie out of his bottle, and we do not know how to put him back.

References

Alchian, A. (1950), 'Uncertainty, evolution and economic theory', *Journal of Political Economy*, **58**, pp. 221–41.

Axelrod, R. (1984), *The Evolution of Cooperation*, New York: Basic Books.

Bicchieri, C. (2006), *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*, Oxford: Oxford University Press.

Bicchieri, C. (2012), *The Grammar of Society: The Nature and Dynamics of Social Norms*, New York: Cambridge University Press.

Binmore, K. (2001), *'A review of Robert Axelrod's: The complexity of cooperation: Agent-based models of competition and collaboration', Journal of Artificial Societies and Social Simulation*, available online at http://jasss.soc.surrey.ac.uk/1/1/review1.html.

Binmore, K. (2005), *Natural Justice*, New York: Oxford University Press.

Binmore, K. (2010), 'Social norms or social preferences?' *Mind and Society*, **9**, pp. 139–58.

Binmore, K. (2020), *Crooked Thinking or Straight Talk? Modernizing Epicurean Scientific Philosophy*, Cham: Springer Nature.

Binmore, K. and Samuelson, L. (1994), 'Equilibrium selection and evolutionary drift', *Review of Economic Studies*, **66**, pp. 363–94.

Friedman, M. (1953), 'The methodology of positive economics', in *Friedman's Essays on Positive Economics*, Chicago: Chicago University Press.

Hindricks, F. and Guala, F. (2015), 'Institutions, rules and equilibria: A unified theory', *Journal of Institutional Economics*, **11**, pp. 459–80 (see also the commentaries in the same issue).

Hume, D. (1739), *A Treatise of Human Nature*, Oxford: Clarendon Press (edited by Selby-Bigge, L. A. Revised by Nidditch, P., 1978.)

Kirman, A. and **Vriend, N.** (2000), 'Learning to be loyal: A study of the Marseille fish market', in Gallegati, G. and Kirman, P. (eds), *Interaction and Market Structure, Lecture Notes in Economics and Mathematical Systems 484*, Berlin: Springer.

Lewis, D. (1969), *Conventions: A Philosophical Study*, Cambridge, MA: Harvard University Press.

Maynard Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.

Pollock, G., Cabrales, A., Rissing, S. and Binmore, K. (2012), 'Suicidal punishment in the ant *Acromyrmex versicolor*', *Evolutionary Ecological Research*, **14**, pp. 1–21.

Schelling, T. (1960), *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.

Tinbergen, N. (1963), 'On aims and methods of ethology', *Zeitschrift fur Tierpsychologie*, **20**, pp. 410–33.

Trivers, R. (1971), 'The evolution of reciprocal altruism', *Quarterly Review of Biology*, **46**, pp. 35–57.

Young, P. (2001), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton, NJ: Princeton University Press.

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press

A. Appendix. David Hume Quotes

This appendix pays tribute to Hume (1739) as the philosopher who first understood the significance of social norms.

A.1. The evolution of conventions

Two men, who pull the oars of a boat, do it by an agreement or convention, though they have never given promises to each other. Nor is the rule concerning the stability of possession the less derived from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. On the contrary, this experience assures us still more, that the sense of interest has become common to all our fellows, and gives us a confidence of the future regularity of their conduct: And it is only on the expectation of this, that our moderation and abstinence are founded. In like manner are languages gradually established by human conventions without any promise. In like manner do gold and silver become the common measures of exchange, and are esteemed sufficient payment for what is of a hundred times their value.

A.2. Reciprocity

I learn to do service to another, without bearing him any real kindness, because I foresee, that he will return my service in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me and others. And accordingly, after I have served him and he is in possession of the advantage arising from my action, he is induced to perform his part, as foreseeing the consequences of his refusal.

A.3. Long-term relationships

Your corn is ripe today, mine will be so tomorrow. 'Tis profitable for us both that I should labour with you today, and that you should aid me tomorrow, I have no kindness for you and know you have as little for me. I will not, therefore, take any pains upon your account; and should I labour with you upon my own account, in expectation of a return, I know I should be disappointed, and that I should in vain depend upon your gratitude. Here, then I leave you to labour alone. You treat me in the same manner. The seasons change and both of us lose our harvest for want of mutual confidence and security.

A.4. Authority as a social norm

Nothing appears more surprising to those who consider human affairs with a philosophical eye, than the ease with which the many are governed by the few, and the implicit submission with which men resign their own sentiments and passions to those of their rulers. When we inquire by what means this wonder is effected, we shall find that, as Force is always on the side of the governed, the governors have nothing to support them but opinion. It is therefore on opinion only that government is founded, and this maxim extends to the most despotic and most military governments as well as to the most free and most popular.

Cite this article: Binmore, K. (2022), 'Thinking about social norms', *National Institute Economic Review*, **259**, pp. 20–30. https://doi.org/10.1017/nie.2022.2

https://doi.org/10.1017/nie.2022.2 Published online by Cambridge University Press